Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2007	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Multimodal Meeting Capture and Understanding with the CALO Meeting Assistant	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Center for the Study of Language and Information, Stanford University Stanford, CA 94305**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**2**	

Multimodal Meeting Capture and Understanding with the CALO Meeting Assistant

Patrick Ehlen[1]
and The DARPA CALO Meeting Assistant Project Team[2]

[1] CSLI, Stanford University, Stanford, CA 94305
ehlen@stanford.edu

Abstract. The CALO Meeting Assistant is a multimodal meeting assistant technology that integrates speech, gestures, and multimodal data collected from multiparty interactions during meetings. Using machine learning and robust discourse processing, it provides a rich, browsable record of a meeting.

Keywords: meeting assistant, machine learning, multimodal integration, CALO, shallow discourse understanding, topic segmentation, implicit supervision, multi-party dialogue

1 Introduction

For nearly as long as people have been holding meetings, they have been thinking up ways to make them more bearable. Recent advances include technologies that integrate the processing of speech, natural language understanding, vision, and multimodal interaction to produce tools that perceive what happens at a meeting, extract salient events, and produce a reliable record. The ICSI Meeting Project [1] sought to produce automated and segmented transcripts from natural, multi-party speech during meetings, while the ISL Smart Meeting Room Task [2] and the M4 and AMI projects [3] instrumented meeting rooms to collect data on behaviors so the interactions of meeting participants could be analyzed to produce flexible records of their activities, while providing a supportive environment for collaboration.

Akin to the latter, the CALO Meeting Assistant collects data about the behaviors of people in meetings, assimilating speech, movement, and note-taking data to create a rich representation of a meeting that can be analyzed and reviewed at many levels. In addition, the CALO Meeting Assistant integrates its observations with those of a larger system of agents, which assesses the meeting data in the context of the ongoing projects and workflow for each of the meeting participants. Thus, our meeting assistant aims to reach beyond an intelligent room that understands only the activities of people in meetings, and attempts to understand their overarching concerns and interpret their behaviors from the perspective of what their meetings mean to them.

[2] The DARPA CALO MA Project is a collaborative effort among researchers at Adapx, CMU, Georgia Tech, MIT, SRI, Stanford University, UC Berkeley, and UC Santa Cruz.

That larger system of agents—developed under the DARPA CALO (Cognitive Assistant that Learns and Organizes) Program—seeks to produce personalized agents that support high-level knowledge workers in carrying out their professional activities. CALO handles a broad range of decision-making tasks that are traditionally resistant to automation, in part by interacting with and learning from its users. It takes initiative on completing routine tasks, and on assisting when the unexpected happens, by maintaining explicit models of ongoing activities and operating environments, adapting and learning "in the wild" while it is used.

The CALO Meeting Assistant Project allows CALO to participate in group discussions—particularly, meetings. Unlike instrumented "intelligent room" meeting projects, it is designed for users in an office environment with a laptop and some small, off-the-shelf peripheral devices (such as headsets, webcams, and digital writing devices) to capture speech, gestures, and handwriting. It aims to be unobtrusive, leveraging cross-training, unsupervised learning, and lightweight supervision captured from normal user interaction (such as reviewing notes, or adding tasks to a to-do list).

The CALO Meeting Assistant captures speech, pen, and other meeting data using an ordinary laptop; produces an automated transcript; segments by topic; and performs shallow discourse understanding to produce a list of probable action items arising from a single, pre-recorded meeting. A *Meeting Rapporteur* provides a meeting summary, and allows participants to review and organize the meeting transcript, audio, notes, action items, and topics—all while harvesting user actions in a feedback loop that supports the meeting assistant's semi-supervised learning process. And potential real-time capabilities under development will allow users to interact with the meeting assistant during an ongoing meeting.

Acknowledgments. This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Contract No. NBCHD030010. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of DARPA or the Department of Interior-National Business Center.

References

1. Janin, A., Ang, J., Bhagat, S., Dhillon, R., Edwards, J., Marcias-Guarasa, J., Morgan, N., Peskin, B., Shriberg, E., Stolcke, A., Wooters, C., and Wrede, B. The ICSI Meeting Project: Resources and Research. In Proc. of ICASSP 2004 Meeting Recognition Workshop (NIST RT-04) (2004)
2. Nijholt, A., op den Akker, R., and Heylen, D.: Meetings and Meeting Modeling in Smart Environments. AI & Society 20 (2005) 202-220
3. Waibel, A., Schultz, T., Bett, M., Denecke, M., Malkin, R., Rogina, I., Stiefelhagen, R., and Yang, J. SMaRT: The Smart Meeting Room Task at ISL. In Proc. of ICASSP (2003) 752-755